A reference database of *Ralstonia solanacearum* egl-mutS haplotypes for global epidemiological surveillance of bacterial wilts

Emmanuel WICKER[1], Carine N'GUESSAN[2], Anne-Claire LE ROUX-NIO[3], Peninna DEBERDT[4], Lakshmi SUJEEUN[1], Pierre RIVIERE[1], Boris VINATZER[5]

[1]CIRAD, UMR PVBMT, Saint Pierre, Reunion; [2]University Khorogo, Côte d'Ivoire; [3]FN3PT, INRA Le Rheu, France; [4]CIRAD, CAEC, Le Lamentin, Martinique; [5]Department PPPWS, Virginia Tech, Blacksburg, VA, USA

What's in your field ?

- Accurate determination of the *Ralstonia solanacearum* Species Complex (RSSC) strains epidemiologically active is a crucial information for designing adequate control strategies
- Despite numerous genomic resources, scientists working on new outbreaks still fail to position their strains in the intraspecific classification, due to the lack of a reference database.
- Using two MLST reference genes (*egl* and *mutS*), we propose a publicly available set of reference sequences allowing to position new strains within the phylotype-clade-sequevar-ST system.

Phylotype	Origin	Clade	*egl*-based sequevars (Fegan & Prior, 2005)	Biovar	Race	Ecotypes
I	Asia (and East-Africa ?)	1	13 14 15 17 18 / 30 31 32 33 34 / 44 45 46 47 48	3	1	Bacterial Wilt Solanaceae
			16	4	4	Ginger
			12	5	5	Mulberry
III	Africa	6	19 20 21 22 23 / 42 43 44 48 49	1		Bacterial Wilt Solanaceae
			29	2T		Brown Rot
IIA	Americas	2	6 24 / 5 35 36 37 38 / 39 41 50 52 53	1	2	Moko
					1	Bacterial Wilt Solanaceae
IIA^T		3	7			
IIB	Americas	4	4A 4SFR	1	2	Moko
			4NPB 51	1		Emerging "NPB"
		5	1 2	2	3	Brown Rot
			3	1	2	Moko
			25 26 27	2T		Brown Rot
		ND	28			
IV	Indonesia	7	8	2T		Bacterial Wilt Solanaceae
			10	BDB		Banana BD
		8	9a 9b	RSY		Sumatra disease of clove
		ND	11	2		

The intraspecific classification of the RSSC based on phylotypes, clades, sequevars, and their correspondance to races, biovars, ecotypes (From Fegan & Prior 2005; Prior & Fegan 2006; Denny & Genin 2012; Wicker et al 2012)

The approach

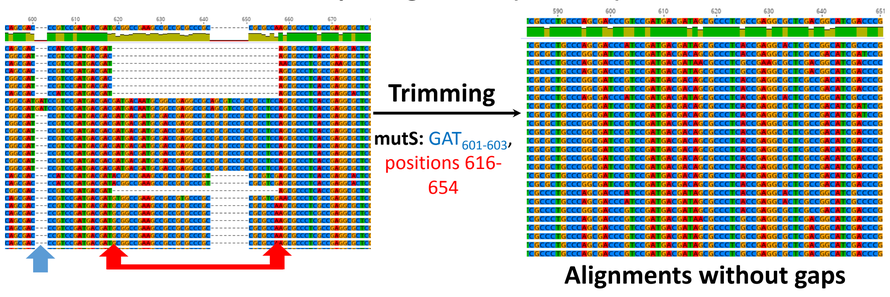
RUN collection

Egl: primers ENDO-F/R (Fegan & Prior, 2005)

mutS: primers mutS.RsF/R (Prior & Fegan, 2005)

Sequences GENBANK

Multiple alignments (CLUSTAL)



Trimming

mutS: GAT$_{601-603}$, positions 616-654

Alignments without gaps

Egl: 666 nt sequences
mutS: 651 nt sequences

- Identification of Sequence Types (DNASP)
- The reference sequences have been deposited in the Plant Associated and Environmental Microbes Database (PAMDB)

http://genome.ppws.vt.edu/cgi-bin/MLST/home.pl

1. An updated view of the global *Ralstonia solanacearum* diversity

Total collection: 1416 strains

- **Egl:** 125 Sequence Types (eST) (207 informative positions /666)
- **mutS**: 120 Sequence Types (mST) (318 informative positions/651)

- Each ST corresponds to a single Phylotype and Clade

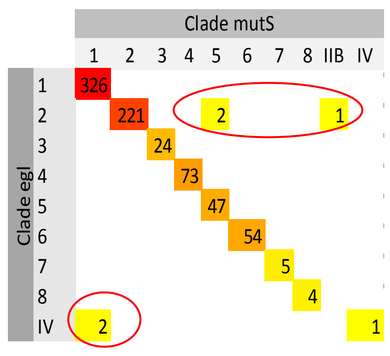
Common collection: 755 strains, typed at both genes

Phylotype	Clade	No strains	*Egl* No ST	*Egl* G/N	*mutS* No ST	*mutS* G/N	Combined (emST) No ST	Combined (emST) G/N
I	1	326	33	0.101	27	0.083	53	0.163
III	6	54	21	0.389	29	0.537	33	0.611
IIA	2	221	19	0.086	16	0.072	33	0.149
	3	24	5	0.208	3	0.125	6	0.250
IIB	4	73	4	0.055	5	0.068	7	0.096
	5	47	9	0.191	12	0.255	14	0.298
IV	7	5	2	0.400	2	0.400	2	0.400
	8	4	3	0.750	3	0.750	4	1.000
	9 (ACH732)	1	1	1.000	1	1.000	1	1.000
TOTAL		755	97	0.128	98	0.130	153	0.203

G/N (number of genotypes/number of strains) measures the actual diversity per clade

- The highest diversity was observed in phylotype III, then IIA/clade 3, then IIB/clade 5, then phylotype I

2. Egl and mutS give congruent phylogenies, but…



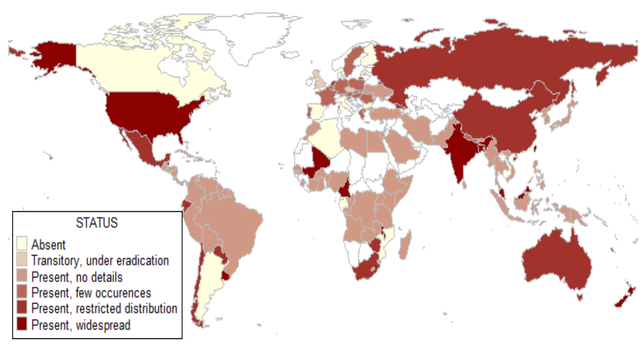
- These incongruences may be due to recombination affecting *egl*

3. The sequevar system is questioned

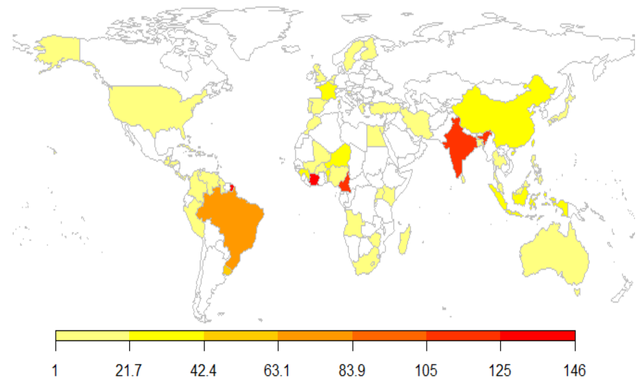
- 14 sequevars differ by less than 1% (6.66 nt) from each other
- 60 eSTs are associated to different sequevars under the 6 nt threshold

eST	Differences in nt under the sequevar threshold: 6	5	4	3	2	1
1	SEQ38	SQ24 SQ38	SQ38 SQ35	SQ39	**SQ39**	
5	SQ13 SQ14 SQ15 SQ16	SQ12 SQ13 SQ14 SQ15 SQ16 SQ34	SQ13 SQ15 SQ45 SQ46 SQ48	SQ13 SQ17 SQ18 SQ31 SQ48 SQ46 SQ47	SQ48 SQ17 SQ18 SQ31 SQ45	**SQ17**
14	SQ13 SQ17	SQ13 SQ46 SQ31 SQ15 SQ12 SQ47	SQ45	SQ14	SQ14 SQ14	**SQ14**
15	SQ45 SQ48 SQ13	SQ14 SQ48	SQ13 SQ46 SQ48	SQ31 SQ14	SQ14 SQ15	**SQ15** SQ16
18	SQ38	SQ38 SQ50	SQ39 SQ35	**SQ35**		
19	SQ14 SQ17	SQ14 SQ16 SQ44 SQ46 SQ48	SQ14 SQ15 SQ14 SQ17 SQ18 SQ34 SQ46 SQ47	SQ13 SQ14 SQ17 SQ18 SQ45 SQ48	SQ17 SQ13 SQ31	**SQ13** **SQ31** **SQ45**

4. A database improving *R.solanacearum* global surveilance



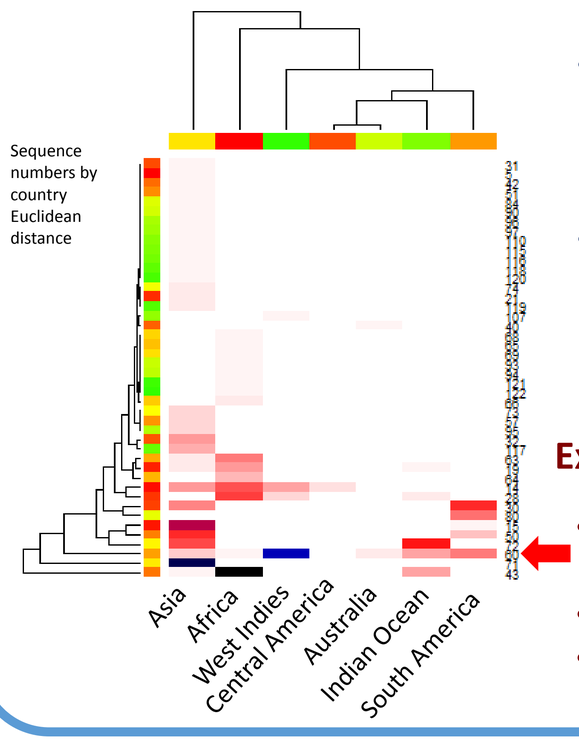
Ralstonia solanacearum distribution map (EPPO,June 2016)



World distribution map of our sequences

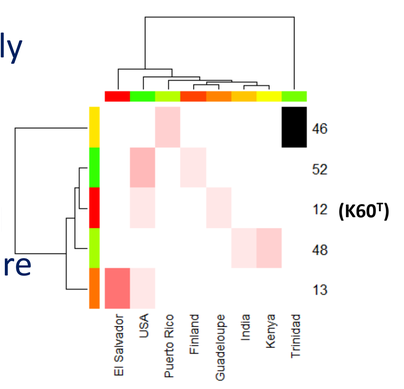
- Sequences available from countries where Rs is declared absent by EPPO: Finland, Haiti, Puerto Rico, Seychelles, Spain, Guinea, Niger, Israel
- 56 « infected » countries with no data, mainly in Africa. **Strong need for surveys in East and Austral Africa, andean latin America, China.**

5. A database to identify pandemic and region-specific STs



Example 1: Within Phylotype IIA-Clade 3

- This clade (formerly sequevar 7) is not limited to South-Eastern USA !
- The Carribean and India-Kenyan STs are original.



Example 2: Within Phylotype I

- The eST060/mST022 (GMI1000) is the most widespread ST, but absent from Africa.
- Asia is the richest region, followed by Africa.
- Africa contains 8 specific STs.

The phylotype-clade-emST system is a good proxy for rapid identification of new lineages, and guiding management strategies. A complete overview of this database is in preparation.

Acknowledgements

Fanny MAILLOT, Claudine BOYER, Amaala MUHAMAD, Edith LALLEMAND-MAMOSA, Marie TERVILLE, Frédéric CHIROLEU